DOCUMENT RESUME

ED 304 473 TM 012 854

AUTHOR Marso, Ronald N.; Pigge, Fred L.

TITLE Standardized and Competency Testing Programs: Classroom Teachers' Needs and Proficiencies as Viewed by Principals and Supervisors.

PUB DATE Oct 88

NOTE 35p.; Paper presented at the Annual Meeting of the Midwestern Educational Research Association (Chicago, IL, October 13-15, 1988). For a related document, see TM 012 681.

PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.

DESCRIPTORS Administrator Attitudes; Beginning Teachers; *Classroom Techniques; Elementary Secondary Education; Minimum Competency Testing; *Principals; *Standardized Tests; *Supervisors; Teacher Evaluation; Teacher Made Tests; *Teaching Skills; Testing Programs; Test Selection; *Test Use

IDENTIFIERS *Teacher Competencies

ABSTRACT

This study involved the assessment of classroom teachers' standardized and mastery program testing needs and proficiencies by a random sample of Ohio principals and supervisors (N=586). A total of 229 supervisors, 313 building principals, and 44 individuals in related supervisory roles was used. The findings from the investigation support the conclusions of other researchers that teachers need more training in testing skills and that pre-service training in testing skills may not be meeting the needs of classroom teachers. Both supervisors and principals rated beginning teachers' proficiencies lower than they rated the teachers' need for these skills in order to perform successfully in the classroom. Supervisors and principals rather consistently agreed with each other about the level of the teachers' testing proficiencies and the level of the needs of the various testing skills to assure teaching success. Neither teachers' mastery nor competency testing program needs or proficiency ratings differed significantly when they were classified by school setting (rural, urban, and suburban); however, both teachers' needs and proficiencies for standardized testing program ratings differed significantly when the teachers and administrators were classified by their grade level responsibilities (elementary, middle, and secondary). Four tables present study rating means. (Author/SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED304473

Standardized and Competency Testing Programs: Classroom Teachers' Needs and Proficiencies As Viewed by Principals and Supervisors

Ronald N. Marso

and

Fred L. Pigge

College of Education and Allied Professions

Bowling Green State University

Bowling Green, Ohio 43403

A paper presented at the annual meeting of the

Midwestern Educational Research Association

Chicago

October 13-15, 1988

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Ronald N. Marso

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Running Head: NEEDS AND PROFICIENCIES

TM012854

Abstract

This study involved the assessment of classroom teachers' standardized and mastery program testing needs and proficiencies by a random sample of Ohio principals and supervisors (N=586). The findings from the investigation lent support to the conclusions of other researchers that teachers need more training in testing skills and that preservice training in testing skills may not be meeting the needs of classroom teachers. Both supervisors and principals rated beginning teachers' testing proficiencies lower than they rated the teachers' need for these skills in order to perform successfully in the classroom. Supervisors and principals rather consistently agreed with one another about the level of the teachers' testing proficiencies and the level of the needs of the various testing skills to assure teaching success. Neither teachers' mastery nor competency testing program needs or proficiency ratings differed significantly when they were classified by employing school setting (rural, urban, and suburban); however, both teachers' needs and proficiencies for standardized testing program ratings differed significantly when the teachers and administrators were classified by their grade level responsibilities (elementary, middle, and secondary).

Standardized and Competency Testing Programs: Classroom Teachers' Needs and Proficiencies As Viewed by Principals and Supervisors

Educational policymakers have advocated increased use of tests in the 1980's as a solution to perceived weaknesses in public education for what some educators have called measurement-driven instruction. This movement in turn has led to a heightened concern about testing and teachers' competence in the use and interpretation of these tests. Even though policymakers appear to have confidence in testing as a major mechanism for reforming our educational system, the general educational community as well as the professional testing community remain skeptical that testing alone is sufficient to meet the task of educational reform (Madaus, 1985). For example, Brandt stated in his introduction to the October (1985) issue of Educational Leadership which was devoted to solutions to the testing problem in the schools that the vast majority of teachers and principals distrust standardized tests, that they are weary of being compared and having children compared in ways they consider destructive, and that they see evidence that the testing tail is increasingly wagging the curriculum dog.

It appears that we have limited knowledge about the actual impact of testing upon day-to-day classroom activities and in particular about what Linn (1983) referred to as the link between

testing and instruction. The exact nature of this "link" between testing and instruction and how it functions in the day-to-day classroom setting appears to be largely unknown for teacher-made tests (Gullickson, 1984) and to be a relatively unclear, weak and indirect linkage for standardized tests (Tyler & Sheldon, 1979).

When teachers were asked to provide self-reports about their understanding of and use of standardized test results in their classroom activities, Stetz and Beck (1981) found that only 59% of their sample of 3000 teachers said that they made "some" to "considerable" use of standardized test results. Similarly, Stiggins and Bridgeford (1985) in a survey of 228 teachers found that only 45% reported being "comfortable" in the use of standardized test scores and that another 35% reported no use of the scores. Further, when this latter sample of classroom teachers was asked to check their specific uses of standardized test scores, just 17% checked diagnosis of student performance, just 25% checked instructional grouping, and just 19% checked reporting to others as their uses for standardized test scores. Similarly, Kuhs, Porter, Floden, Freeman, Schmidt, and Schwille (1985) reported that standardized test scores had little influence on the daily instruction of their sample of elementary grade teachers.

Salmon-Cox (1981) used extensive interviews to determine the views of 68 elementary teachers regarding the roles and uses of

standardized achievement tests in the public schools. She found that these teachers perceived the primary role of schools to be a "socializing" function; whereas the "basic skills" function (and related formal measurement of skills) was perceived to be subordinate to the socializing function. Secondly, these teachers reported that "observations" and "teacher-pupil social interactions" were their primary means of assessing student performance and that the results of these assessment strategies were of much more significance than the results of standardized achievement tests in influencing their classroom instruction. These teachers described standardized achievement tests as being too narrow in curricular scope and as providing feedback too delayed to be of practical value in day-to-day classroom decision-making. Further, these teachers described standardized aptitude tests as primarily fulfilling a "gatekeeping" function in assigning students to special education programs and as not primarily serving a classroom instructional function. The role of standardized achievement tests in classroom instruction, however, was perceived by these teachers as being broader than that of the aptitude tests. Nevertheless, the role of the standardized achievement tests was described as being limited to: a) a supplementary or confirmatory function relative to the information gathered through observations, b) a guide or check on pupils' curricular progress perhaps leading to a decision to

spend more time on some curriculum area, and c) an adjunctive check on previously made student grouping decisions. In light of these teacher perceptions of the functions of tests and how these elementary teachers described their use of test scores in the classroom, Salmon-Cox concluded that the frequently expressed concern of teachers' misuse or over-use of test scores appears to be unwarranted. For example, she found that students' low standardized test scores tend to be discounted by classroom teachers while students' high standardized test scores act as a "red flag" to teachers signaling for an upward reassessment of their expectation of students. She also found that these teachers did not allow test score knowledge to result in the "labeling" of their students.

The conclusion of Salmon-Cox, that standardized test scores have little effect upon teacher ratings of students and that when an influence does occur it results in the teacher's upward reassessment of a student, is supported by an experiment reported by Airasian, Madaus, Kellaghan, and Pedulla (1977). These researchers studied the effects of the presence of standardized achievement test scores on second grade teachers' ratings of students. They found that the presence of test scores influenced changes in teacher ratings of their students in only 7% to 10% of the cases and that all changes resulted in increased ratings of student performance. In other words, this sample of teachers

also discounted lower than expected standardized achievement scores but valued and acted upon the higher than expected scores.

Relative to educational administrators' use of standardized test scores, Sproull and Zubrow (1981) reported the results of extensive interviews with 58 administrators in 18 school systems using much the same interview format as used in the previously reported Salmon-Cox study. They interviewed administrators in charge of school testing programs, superintendents, directors of curriculum and instruction, and directors of pupil personnel services. These administrators described the general purpose of standardized testing as being for a) student diagnosis and placement, b) program evaluation, and c) reporting student and school progress to others. They reported infrequent use of individual students' standardized test scores while describing curriculum checks and comparisons of current class performance with national or local norms as their predominant uses of the results of standardized tests. They also indicated that they did not act on class or group scores alone, but rather would raise questions relative to discrepant scores with responsible staff after which a collective decision would be made. These administrators further stated that they believed that the primary benefits of a school standardized testing program accrued to others, namely, those closely tied to the instructional activities. Conversely, the teachers interviewed by Salmon-Cox

reported that they felt that the primary benefits of their school's testing program accrued not to themselves but to others. Thus, it would appear that neither teachers nor administrators feel that school standardized testing programs operate primarily for their benefit.

Regarding teacher training for and teacher proficiency in the interpretation and use of standardized tests, the educational literature commonly appears to convey the assumption that both are inadequate (Mehrens & Lehman, 1987). Illustrative of education literature wherein this assumption is made is Lambert's (1980-81) nation-wide survey of officials of teachers associations, of chairpersons of state legislative educational committees, and of deans of the three largest teacher training institutions in each state. Based upon these self-reports of individuals who were basically removed from day-to-day classroom practices, Lambert concluded that: a) teachers' attitudes toward standardized tests are negative, b) teachers need to learn more about tests and their use, and c) teachers need to be cautioned against over-reliance on tests. Despite what appeared to be almost a consensus among the three groups of respondents that classroom teachers need further training in testing skills, Lambert found that approximately 25% of the responding deans indicated that their institutions did not require and did not intend to require an evaluation or tests and measurements course

to be taken by their teacher candidates. A few studies that gathered data directly from practicing teachers support the assumption that teacher training and teachers' proficiencies in testing are inadequate. For example, Yeh (1978) and Fleming (1979) reported that inservice teachers desire more training in measurement skills, and the results of still other studies suggest that teachers' testing needs are not being well met by either preservice or inservice training (Gullickson, 1986; Gullickson & Ellwein, 1985; Wanous & Mehrens, 1981).

Purpose

The overall purpose of this study was to gather and analyze data related to teachers' need for and their proficiency in various competencies related to school standardized and competency testing programs. More specifically, this study was designed to solicit principals' and supervisors' assessments of their teachers' need for various testing competencies and also to solicit principals' and supervisors' assessments of their typical beginning teachers' levels of proficiency in these same competency areas.

The following four hypotheses guided the investigation:

1) The supervisors and principals will indicate that their beginning teachers are adequately trained relative to the competencies needed for standardized and competency testing programs in their schools. Specifically, it

was hypothesized that the administrators' ratings of the teachers' need for the standardized and competency (mastery) testing competencies would not differ significantly from their ratings of beginning teachers' proficiencies in these same competency areas.

2) The supervisors and principals with different grade level and school type assignments will not differ in their ratings of classroom teachers' testing needs and related beginning teachers' proficiencies. Specifically, it was hypothesized that the ratings of classroom teachers' testing needs and beginning teachers' proficiencies would not differ significantly when administrators' ratings were classified by: a) elementary as compared to middle and high school grade assignments, and b) rural as compared to urban and suburban school assignments.

3) The supervisors and principals will rate their beginning teachers' competencies related to standardized and competency testing programs as high or higher than they will rate their beginning teachers' competencies in: a) knowledge of their subject area, b) their other professional education competencies, and c) their overall competencies as educators.

4) The supervisors and the principals will be in accord regarding their perceptions of the relative levels of teachers' needs and proficiencies associated with standardized and competency testing programs. Specifically, it was hypothesized that the supervisors' ratings when compared to the principals' ratings would not differ significantly for their: a) teachers' testing competency needs, and b) beginning teachers' proficiencies in the testing competencies.

Method

An instrument designed to assess classroom teachers' testing needs and proficiencies was constructed and mailed to a stratified random sample of teacher supervisors and building principals in Ohio during the winter of 1986. The names and school addresses of these administrators were randomly selected from the state directory of schools with the directory classifications of type of school system (city, exempted village, and county local), administrative assignment (principal and supervisor), and school grade level (elementary, middle, and secondary) used as strata. A total of 800 assessment instruments were mailed from which 586 (73%) usable responses were obtained after two follow-up contacts of nonrespondents. A total of 229 supervisors, 313 building principals, and 44 individuals in

related supervisory roles (coordinators of curriculum or instruction, etc.) returned usable assessment forms.

The part of the assessment instrument related to the present paper consisted of a 15-item listing of competencies related to the interpretation and use of standardized (purchased) tests and minimum competency (mastery) tests. These items were selected from a longer list of researcher constructed items and then reviewed for appropriateness by a team of five professors primarily responsible for the instruction of the tests and measurements course required of preservice teachers at Bowling Green State University. The items were grouped into two sections of the assessment instrument with nine items identified as purchased test related competencies and six items identified as competency or mastery testing program competencies. Two five-point ('5' as high and '1' as low) Likert-type response scales were provided for each competency item and were identified as: "need of this competency to be a successful teacher in your school" and "average proficiency of your typical beginning teachers in this competency." The raters were directed to disregard specialized area (art, music, shop, etc.) and special education teachers and to respond relative to the needs and proficiencies of elementary grade and the various subject area teachers. Each respondent was also asked to indicate the nature of his/her school(s) assignment (rural, urban, or suburban) and

the grade level range of his/her primary assignment (elementary, middle grades, secondary, K-12 grades, or other). Those respondents placing themselves in the "other" or "K-12" grade level categories were excluded from the later analyses related to grade level assignment. Additionally, the respondents were asked to rate the proficiency of their typical beginning teachers' testing and evaluation skills as compared to their proficiency in three other skills areas on Likert-type five-point scales ('1' much below average to '5' well above average). These three other skills areas were identified as: the beginning teachers' subject area knowledge, the beginning teachers' knowledge and skills in other professional education competencies (planning, discipline, etc.), and the beginning teachers' overall skills as educators.

Results

Hypothesis One: Teachers' Level of Need Versus Their Proficiency

A t test of the difference between dependent means was used to analyze the combined supervisors' and principals' ratings (N=586) of their classroom teachers' need for as compared to their typical beginning teachers' proficiency in each identified competency. These t test comparisons were completed between the need rating mean and the proficiency rating mean for each of the nine purchased test competency items and for each of the six competency testing program items; also, t-ratios were completed for both sets of combined items (need as well as proficiency) for

the standardized as well as the mastery sections of the instrument.

The t test analysis procedures resulted in the rejection of hypothesis one as significant differences ($p < .0005$) between the need and proficiency rating means were noted for each of the 15 competency items as well as for the combined totals. Descriptions of the items on the assessment instrument, need and proficiency rating means, t-ratios, and other data related to these analyses are presented in Table 1. For each of the 15 competency items the combined group of respondents (supervisors and principals) rated the classroom teachers' mean need for the competency significantly higher than they rated their typical beginning teachers' average level of proficiency in that same competency area. This would suggest that the administrators felt that their typical beginning teachers' testing proficiencies were inadequate to meet the needs of their jobs. The need and proficiency rating means for the combined items for the purchased test competencies were, respectively, 3.80 and 2.72 ($t = 30.16$, $p < .0005$) and for the mastery testing items the need and proficiency means, respectively, were 4.22 and 2.82 ($t = 20.71$, $p < .0005$).

To better identify which beginning teachers' competencies the respondents reported as being most deficient in proficiency relative to need for classroom success, a discrepancy index was

calculated for each item (need rating mean minus proficiency rating mean); each item was then ranked over the total 15 items relative to the magnitude of this discrepancy index (see Table 1). The mastery program testing items revealed higher discrepancy values (i.e., lower relative teacher proficiencies) than did the competency items related to the purchased or standardized testing program items. These larger discrepancies on the mastery testing items appeared to result primarily from a higher need rating assigned to these competencies by the administrative raters rather than from relatively lower proficiency ratings of the teachers (i.e., the total need mean for the purchased test competencies was 3.80, whereas the comparable total need mean for the mastery testing items was 4.22). The items with the three largest discrepancy scores for the purchased testing program were: using scores to identify student deficiencies; reporting scores to students and parents; and use of school test scores to assess day-to-day curriculum. The three items with the largest discrepancy scores for the mastery testing area were: using scores in making instructional decisions for individual pupils, writing mastery tests for individualized instructional plans, and identifying basic skills and stating objectives in mastery units.

Hypothesis Two: Needs and Proficiencies by Grade and School

A series of one-way ANOVA procedures were used to analyze the ratings of the administrators when they were classified by their grade level assignment (elementary, middle school, and secondary school) and their school assignment (rural, urban, and suburban). For each classification, a separate ANOVA analysis was completed for each of the 15 need and proficiency items and for the two sets of total means.

The ANOVA analyses related to hypothesis two revealed that there were no significant mean differences between urban, rural, and suburban administrator's ratings of teachers' needs or proficiencies for either the purchased test or mastery test competencies. However, when the administrators' ratings were classified by grade level, significant mean differences were identified among the ratings of the teachers' purchased test competency needs and among the ratings of the teachers' purchased test proficiencies as presented on Table 2 but not among the ratings of teachers' mastery test competency needs nor among the ratings of teachers' mastery test.

The elementary administrators rated both the classroom teachers' need for and their beginning teachers' proficiencies on the purchased test competencies higher than did the middle or secondary administrators. The average combined need rating means for the nine items were 4.04 for elementary, 3.77 for middle

grade, and 3.53 for secondary administrators ($F = 23.07$, $p < .0005$); similarly, the combined teachers' proficiency rating means for these three respective groups were: 2.85, 2.72, and 2.65 ($F = 4.31$, $p < .014$). As shown on Table 2, significant mean differences among the administrators' need ratings when grouped by grade level were noted on each of the nine purchased test competencies ($p < .01$); whereas significant mean differences ($p < .05$) were found between the three groups of administrators' rating of beginning teachers' proficiencies on five of the nine purchased test competencies.

The total need means for the purchased test competencies when raters were classified by type of school assignment were: rural 3.74, urban 3.81, and suburban school raters 3.80 ($F = .54$, $p < .58$). The total need means for the mastery competency section were: rural 4.23, urban 4.09, and suburban school raters 4.19 ($F = .18$, $p = .40$). The total proficiency means for the purchased and mastery sections, respectively, were: rural 2.74, urban 2.78, and suburban 2.67 ($F = 1.09$, $p = .34$), and rural 2.78, urban 2.98, and suburban 2.74 ($F = 1.45$, $p = .24$). This particular set of data is not reported in tabular form.

The failure to identify significant mean differences between administrators' ratings when classified by type of school assignment (rural, urban, and suburban) but the identification of significant differences between the administrators' ratings when

classified by the grade assignment classification would suggest that beginning teachers were seen as having similar levels of testing needs and proficiencies in different types of schools; however, the elementary, middle, and secondary administrators perceived classroom teachers' needs for purchased testing (but not mastery testing) competencies and beginning teachers' proficiencies in these competencies differently. Both teachers' needs and teachers' proficiencies in competencies related to purchased or standardized tests were perceived as being lower at the upper grade levels than at the elementary or middle grades.

Hypothesis Three: Testing Versus Other Professional Competencies

As indicated before, the administrative respondents were requested to make overall comparative assessments of the proficiency of their typical beginning teachers in tests and evaluation competencies relative to these three areas: knowledge of their subject fields, their other professional education skills (such as planning, discipline, etc.), and their overall skills as educators.

The responses to this section of the questionnaire were analyzed by grade level and type of school assignment for the total group of respondents and for supervisor ratings as compared to principal ratings. When the combined group of administrators (principals and supervisors) was classified by grade assignment (elementary, middle, or secondary schools) and by type of school

(rural, urban, or suburban), no significant mean differences for any of the three comparisons were found among either the grade level or type of school classifications. The principals' as compared to the supervisors' rating means, however, were significantly different for each of the three items. As can be seen in Table 3, for each of the three items the supervisors' rating mean of the beginning teachers' competencies was significantly lower than the principals' rating mean.

Hypothesis three was rejected as eight of the nine item rating means (see Table 3) for the principal, supervisor, and the total group of respondents were below the scale average (below 3.0). Thus, it is evident that these principals and supervisors generally perceived beginning teachers as being less proficient in testing and evaluation skills as compared to their proficiency in other professional skills.

Hypothesis Four: Comparison of Principal and Supervisor Ratings

A series of independent t tests were used to determine whether or not the supervisors and principals differed significantly in their ratings of teachers' classroom testing needs and beginning teachers' proficiencies on each of the 15 competencies (and totals) constituting the mastery and purchased tests sections of the assessment instrument. The results of these analyses revealed several differences between

the principals and supervisors as reported on Table 4 and thus led to the rejection of the fourth null hypothesis.

The comparisons of principals' ratings with supervisors' ratings of beginning teachers' needs and proficiencies for the 15 competencies resulted in the identification of seven significant individual competency item mean differences ($p < .05$) among the 30 t-ratio comparisons. The finding of seven differences among the 30 comparisons, although not overwhelming, is almost five times the number that would be expected by chance alone with alpha set at $p < .05$. Even though some differences were noted for individual items, lack of significant mean differences on the combined items means for the two sections suggests some agreement between the two groups about both teachers' testing proficiencies and teachers' testing needs. This agreement between the two groups of raters is more evident among the mastery test competencies where just one of the 12 item comparisons resulted in a significant difference.

Each of the five "need" competency items revealing a significant mean difference between the two groups of raters was rated higher by the supervisors as compared to the principals. Further, this pattern of higher supervisors' need ratings than principals' was the general pattern across all 15 competency items.

The series of comparisons between the principals' and the supervisors' mean ratings of beginning teachers' proficiencies in these 15 competency areas revealed only two significant differences ($p < .05$). The overall pattern of differences, although not significant, suggests higher principal ratings of beginning teacher testing and evaluation proficiencies as compared to the supervisors ratings (higher numerical principals' ratings on all but two competencies).

If one examines the principals' and supervisors' responses to the mastery testing competencies separate from the purchased test competencies, it can be concluded that no significant differences existed between the principals' and supervisors' ratings as the one observed difference among 14 comparisons could be expected by chance alone. When, however, the two sections are examined together, one notes that the pattern of rating means (e.g., the pattern of supervisors providing higher need ratings but lower proficiency ratings) is very consistent over both the mastery and purchased tests. This trend would certainly seem to indicate that supervisors and principals rate teachers' needs and proficiencies differently.

Even though the principals tended to rate the beginning teachers' proficiencies higher and classroom needs lower than did the supervisors as presented on Table 4, it is evident from examining the relative ordering of the item rating means that the

two groups of raters were in very high agreement about the relative levels of classroom teachers' and proficiencies (e.g., which skills were higher or lower among the 15 competencies). To assess the extent of this relative agreement between the principals' and supervisors' proficiency rating means and between their needs rating means over the 15 competencies, both sets of 15 item means (needs means and proficiencies means) were rank ordered and Spearman Rho coefficients of .96 and .93 respectively, were calculated. The magnitude of these coefficients confirms that the supervisors' and principals' ratings of both teachers' needs and teachers' proficiencies suggest an agreement between the two rater groups in the ordering of the importance of the various testing needs and in the ordering the various teachers' proficiencies.

In respect to the nature of the competency items revealing significant differences between the two groups of raters, the supervisors rated just two teachers' proficiencies significantly lower than did the principals: interpreting standard scores, percentile ranks, grade equivalents, etc. and use of test scores to assess day to day curriculum. Conversely, the supervisors rated five competencies significantly higher than did the principals on the needs scale: using scores to identify student deficiencies, interpreting class and individual profiles and charts, use of test scores to assess day to day curriculum,

preparation of students for taking standardized tests, and interpreting criterion-referenced checklists on objectives based tests. Only on a single competency item, use of test scores to assess day-to-day curriculum, did the two groups rate both need and proficiency significantly different.

Summary and Discussion

Review of Findings

The administrators (principals and supervisors as a single group) rated classroom teachers' needs higher than they rated their typical beginning teachers' proficiencies in competencies relative to purchased tests and competency testing programs. This difference suggests that beginning teachers' proficiencies are below the classroom instructional demands as viewed by these administrators.

The administrators assigned to elementary, middle, and secondary level schools differed significantly in their ratings of both the classroom teachers' needs and beginning teachers' proficiencies associated with purchased testing program competencies. The secondary level administrators tended to rate both classroom instructional needs and teachers' proficiencies related to purchased testing competencies lower than did the administrators assigned to the elementary and middle grade levels. Conversely, however, the administrators' ratings of the competency (mastery) testing program items did not differ by

grade level assignment or by type of school assignment (rural, urban, and suburban) and also the administrators' ratings of the teachers' standardized testing item needs and proficiencies did not differ by type of school assignment.

The administrators did not rate typical beginning teachers' proficiencies in testing and evaluation skills as high or higher than they rated the same teachers' knowledge in their subject areas, teachers' other professional competencies (e.g., classroom discipline, etc.), or teachers' overall proficiencies as educators. It was also found that the principals' rating means differed from those of the supervisors' rating means on seven of the 30 teachers' testing proficiency and need means. The supervisors generally rated needs higher and teachers' proficiencies lower than did the principals.

Discussion

The data obtained from this investigation would tend to support Lambert's (1980-81) conclusion that teachers need more training in the field of testing and evaluation and the conclusion of Gullickson (1986) and others that preservice training in testing and evaluation may not be meeting the needs of classroom teachers. Further, this data suggests that beginning teachers may be less proficient in testing and evaluation skills as compared to their other professional skills.

The significant differences found between administrators' ratings when classified by grade level responsibilities indicate that they consider standardized achievement tests to be less needed in the upper grades as compared to the lower grades and they perceive upper grade level teachers to be less proficient in working with purchased achievement tests than are teachers in lower grades. Both of these findings may reflect nothing more than the practice of administrators and teachers in the elementary grades placing more emphasis upon pupils' basic skill development and related frequent standardized achievement testing than do secondary level administrators and teachers.

The lack of differences found between administrators' ratings when grouped by type of school assignment indicate that neither classroom instructional demands nor the proficiencies of typical beginning teachers related to standardized and competency testing programs vary among rural, urban, and suburban schools. This finding would suggest that teachers' basic testing skills and teachers' needs of these skills do not vary by school setting and would tend to support the contention by university educators that the knowledge and skills addressed in preservice tests and measurements training are sufficiently generic to be applicable to various school settings.

The administrators tended to rate nearly all classroom instructional needs for the competency (mastery) testing progra

higher than the teachers' classroom instructional needs for standardized testing. Perhaps this greater emphasis upon classroom instructional needs in the competency programs simply reflects its closer integration of testing and instruction. Similarly, the competency items more closely associated with the actual instructional process within the purchased tests section of the assessment instrument tended to be rated as having higher needs (e.g., use of scores to assess day-to-day curriculum, using scores in assessing reading and related deficiencies, and counseling students on the basis of test results) as compared to the ratings of the need for the less instructionally related items (e.g., administration of standardized tests and interpreting reliability and validity of standardized tests). Likewise, the discrepancies between need ratings and proficiency ratings were greater for most of the competency testing program items than for the standardized testing items, and were greater for the competency items more closely associated with the instructional process than for those competency items less closely associated with the instructional process within the standardized tests section.

It was also found that the larger discrepancies identified between classroom needs and teachers' proficiencies more frequently resulted from administrators' higher classroom instructional need ratings for certain competencies rather than

from administrators' relatively lower ratings of beginning teachers' proficiencies for a particular competency. As a consequence, it may behoove preservice and inservice teacher trainers to simultaneously consider teachers' needs for various testing competencies as well as teachers' relative proficiency levels in these competencies in their instructional planning.

References

Airasian, P. W., Madaus, G. F., Kellaghan, T., & Pedulla, J. J. (1977). Proportion and direction of teacher rating changes of pupils' progress attributable to standardized test information. Journal of Educational Psychology, 69, 702-709.

Brandt, R. (1985). The search for solutions. Educational Leadership, 43, 3.

Fleming, M. (1979). Classroom measurement needs in 1980's. Paper presented at the annual meeting of the National Council on Measurement in Education, San Francisco.

Gullickson, A. R. (1984). Teacher perspectives of their instructional use of tests. Journal of Educational Research, 77, 244-248.

Gullickson, A. R. (1986). Teacher education and teacher-perceived needs in educational measurement and evaluation. Journal of Educational Measurement, 23, 347-354.

Gullickson, A. R., & Ellwein, M. C. (1985). Post hoc analysis of teacher-made tests: The goodness-of-fit between prescription and practice. Educational Measurement: Issues and Practice, Spring, 15-18.

Kuhs, T., Porter, A., Floden, R., Freeman, D., Schmidt, W., & Schwille, J. (1985). Differences among teachers in their use of curriculum-embedded tests. The Elementary School Journal, 86.

Lambert, B. R. (1980-81). Teacher attitudes on testing: A multiple perspective. College Board Review, Winter, 13-14 and 29-30.

Linn, R. L. (1983). Testing and instruction: Links and distinctions. Journal of Educational Measurement, 20, 179-189.

Madaus, G. F. (1985). Public policy and the testing profession: You've never had it so good? Educational Measurement Issues and Practices, 4, 5-11.

Mehrens, W. A., & Lehmann, I. J. (1987). Using standardized tests in education. White Plains, NY: Longman, Inc.

Salmon-Cox, L. (1981, May). Teachers and standardized achievement tests: What's really happening. Phi Delta Kappan, 631-634.

Sproull, L., & Zubrow, D. (1981, May). Standardized testing from the administrative perspective. Phi Delta Kappan, 628-631.

Stetz, F. P., & Beck, M. D. (1981). Attitudes toward standardized tests: Students, teachers, and measurement specialists. Measurement in Education, 12, 1-11.

Stiggens, R. J., & Bridgeford, N. J. (1985). The ecology of classroom assessment. Journal of Educational Measurement, 22, 271-286.

Tyler, R. W., & Sheldon, H. W. (1979). Testing, teaching, and learning. Report of a Conference on Research on Testing. Washington, D.C.: National Institute of Education.

Wanous, D. S., & Mehrens, W. A. (1981). Helping teachers use information: The data box approach. NCME Measurement in Education (special reports series), 12, 1-10.

Yeh, J. P. (1978). Test use in schools. Washington, D.C.: U.S. Department of Health, Education and Welfare, National Institute of Education.

5/2

Table 1

Administrators' Rating Means of the Need and Proficiency of Beginning Teachers' Testing Related Competencies with Discrepancy Between Means, Discrepancy Rank, and t Ratio

Purchased Tests & Scores	N*	Rating Means: Need	Rating Means: Proficiency	Mean Discrepancy	Rank Order of Discrepancy	t	p**
1. Admin. standardized tests	520	3.56	2.95	-.61	15	13.23	.000
2. Interpret rel. & validity	517	3.36	2.51	-.85	14	17.23	.000
3. Interpret types of scores	520	3.72	2.58	-1.14	10.5	22.17	.000
4. Report scores to student and parents	520	4.07	2.84	-1.23	8	25.89	.000
5. Counsel students from scores	521	3.84	2.70	-1.14	10.5	24.46	.000
6. Scores to identify deficiencies	523	4.19	2.82	-1.37	5	30.00	.000
7. Interpret profiles/charts	523	3.80	2.69	-1.11	12	23.84	.000
8. Scores to assess daily curriculum	522	3.80	2.58	-1.22	9	24.03	.000
9. Prepare students for tests	521	3.82	2.73	-1.10	13	22.45	.000
Combined Items	498	3.80	2.72	-1.08	--	30.16	.000
Mastery Testing Programs							
1. Scores direct individual instruction	174	4.45	2.87	-1.58	1	20.56	.000
2. Interpret criterion scores	173	4.18	2.77	-1.41	4	17.98	.000
3. Write mastery tests	172	4.12	2.66	-1.46	2	17.00	.000
4. Prepare students minimum skills	173	4.22	2.91	-1.31	6	16.73	.000
5. State skills and objectives	173	4.36	2.93	-1.43	3	19.03	.000
6. Assessment of mainstreamed students	170	4.02	2.76	-1.26	7	15.97	.000
Combined Items	162	4.22	2.82	-1.40	--	20.71	.000

*The number of administrators' responses for the mastery (competency) testing program section was much smaller as the respondents were directed to respond to this section only if they had such programs operating in their school(s).

**p's of .000 are to be interpreted as less than .0005.

Table 2

Administrators' Needs and Proficiencies Rating Means for Beginning Elementary, Middle, and Secondary Teachers' Purchased Testing Competencies and Related F Values*

Competencies	Elem.	Middle	Secondary	F	p**
	Need Rating Means				
1. Admin. standardized tests	3.98	3.50	3.17	30.52	.000
2. Interpret rel. & validity	3.59	3.37	3.19	6.03	.003
3. Interpret types of scores	3.99	3.79	3.43	16.62	.000
4. Report scores to student and parents	4.33	4.00	3.81	13.85	.000
5. Counsel students from scores	3.95	3.91	3.61	5.94	.003
6. Scores to identify deficiencies	4.38	4.20	3.96	11.58	.000
7. Interpret profiles/charts	4.11	3.67	3.48	21.04	.000
8. Scores to assess daily curriculum	3.99	3.80	3.54	9.43	.000
9. Prepare students for tests	4.00	3.79	3.63	7.00	.001
Total Combined Items	4.04	3.77	3.53	23.07	.000
Competencies	Proficiency Rating Means				
1. Admin. standardized tests	3.19	2.86	2.75	12.22	.000
2. Interpret rel. & validity	2.56	2.56	2.47	.52	.597
3. Interpret types of scores	2.68	2.78	2.62	.86	.425
4. Report scores to student and parents	2.98	2.86	2.79	1.95	.144
5. Counsel students from scores	2.87	2.72	2.65	2.02	.134
6. Scores to identify deficiencies	3.13	2.86	2.69	4.98	.007
7. Interpret profiles/charts	2.87	2.74	2.61	4.24	.015
8. Scores to assess daily curriculum	2.72	2.47	2.60	3.19	.042
9. Prepare students for tests	2.92	2.71	2.61	6.74	.001
Total Combined Items	2.85	2.72	2.65	4.31	.014

N = 423, elementary 162, middle school 76, and secondary school 185.

*Ratings on Mastery Testing Program items did not differ significantly by grade level assignment of the administrators for either the needs or proficiency rating.

**p's of .000 are to be interpreted as less than .0005.

Table 3

Principals' and Supervisors' Rating Means for Beginning Teachers' Testing and Nontesting Competencies

Relative Proficiency Rating Items*	$\bar{X}$ Principal	$\bar{X}$ Supervisor	$\bar{X}$ Combined	t**	p
1. Relative to knowledge of their subject areas, beginning teachers' test and evaluation competencies are...	3.03	2.87	2.95	2.47	.014
2. Relative to their other professional education competencies, such as planning, discipline, etc., beginning teachers' test and evaluation competencies are...	2.96	2.81	2.89	2.34	.020
3. Relative to their overall competencies as educators, beginning teachers' test and evaluation competencies are...	2.93	2.73	2.84	3.34	.001

*Ratings were recorded via a five point Likert-type scale, 5 (well above average), 4 (somewhat above average), 3 (average), 2 (somewhat below average), and 1 (much below average)

**Ratios for t comparisons between the principals' and supervisors' rating means

Table 4

Supervisors' and Principals' Rating Means for Beginning Teachers' Testing Competency Needs and Proficiencies

	Needs				Proficiencies			
Purchased Test Competencies	Prin.	Super.	t	p	Prin.	Super.	t	p
1. Admin. standardized tests	3.49	3.64	-1.66	.10	2.90	3.01	-1.49	.14
2. Interpret rel. & validity	3.33	3.35	-.09	.85	2.54	2.46	1.10	.27
3. Interpret types of scores	3.67	3.75	-.96	.34	2.65	2.47	2.54	.01
4. Report scores to student and parents	4.02	4.08	-.81	.42	2.87	2.80	.96	.34
5. Counsel students from scores	3.78	3.88	-1.23	.22	2.73	2.65	1.09	.27
6. Scores to identify deficiencies	4.11	4.25	-1.94	.05	2.85	2.77	1.15	.25
7. Interpret profiles/charts	3.67	3.92	-3.13	.00	2.75	2.62	1.91	.06
8. Scores to assess daily curriculum	3.69	3.89	-2.47	.01	2.65	2.48	2.38	.02
9. Prepare students for tests	3.73	3.90	-2.26	.02	2.78	2.67	1.49	.14
Combined Items	3.73	3.85	-1.90	.06	2.76	2.66	1.68	.09
Mastery Testing Competencies								
1. Scores direct individual instruction	4.37	4.53	-1.51	.13	2.92	2.83	.76	.45
2. Interpret criterion scores	4.07	4.31	-2.02	.05	2.82	2.73	.70	.49
3. Write mastery tests	4.09	4.11	-.15	.88	2.71	2.61	.80	.42
4. Prepare students minimum skills	4.22	4.20	.21	.83	2.89	2.93	-.40	.69
5. State skills and objectives	4.30	4.40	-.91	.36	2.95	2.91	.39	.69
6. Assessment of mainstreamed students	3.98	4.03	-.43	.67	2.84	2.68	1.29	.20
Combined Items	4.17	4.25	-.81	.42	2.85	2.79	.60	.55

N = 542, principals = 313, supervisors = 229

5/2